MEASURING TRADE-OFFS AMONG CRITERIA IN A BALANCED SCORECARD FRAMEWORK: POSSIBLE CONTRIBUTIONS FROM THE MULTIPLE CRITERIA DECISION ANALYSIS RESEARCH FIELD

Fernando A. F. Ferreira

*School of Management and Technology, Polytechnic Institute of Santarém,
Complexo Andaluz, Apartado 295, 2001-904 Santarém, Portugal
Fogelman College of Business and Economics, University of Memphis,
Memphis, TN 38152-3120, USA
E-mail: fernando.ferreira@esg.ipsantarem.pt; fernando.ferreira@memphis.edu*

Abstract. Claimed by the Harvard Business Review as one of the main innovations of recent decades in management systems, the Balanced Scorecard (BSC) has been widely studied and applied in different management contexts. However, despite recent progress and its undeniable merit, the BSC has its own shortcomings. As reported in the literature, it seems generally agreed that the way compensations between criteria within a BSC framework are calculated remains an open issue. Thus, one of the contributions of this study is to augment the theoretical discussion on the potentialities of the multiple criteria decision analysis (MCDA) approach to bring simplicity and transparency to the calculation of compensations (*i.e.* trade-offs) among evaluation criteria within a BSC framework. It seems important to underline that this paper builds on previous work and is not empirical research. However, it aims to extend the discussion to other (new) measurement contexts. We believe that the development and promotion of the integrated use of MCDA tools in BSC applications to measure trade-offs among evaluation criteria in the BSC framework can be a key factor in ensuring more successful implementations and uses of the BSC within firms.

Keywords: Balanced Scorecard (BSC), decision making, framework, MCDA, multiple criteria, trade-offs.

Reference to this paper should be made as follows: Ferreira, F. A. F. 2013. Measuring trade-offs among criteria in a Balanced Scorecard framework: possible contributions from the multiple criteria decision analysis research field, *Journal of Business Economics and Management* 14(3): 433–447.

JEL Classification: C02, C18, L25, M21.

1. Introduction

Present-day economic thinking assumes that performance evaluation is a key element for the promotion of improvement initiatives (Urbonavičius, Ivanauskas 2005; Strand-skov 2006; Zinkevičiūte 2007; Santos *et al.* 2008; Acar, Zehir 2010; Curado, Manica

2010). As stated by Santos *et al.* (2008), *"formal or informal performance measurement is common practice in most organizations and it is well established that this plays a critical role in signalling the level of success in achieving objectives and identifying where improvement efforts are required"*. From this assumption, remarkable progress has occurred in recent years in the development of performance measurement frameworks, where the Balanced Scorecard (BSC) (Kaplan, Norton 1992) is one of the best known examples. In fact, as claimed by the Harvard Business Review, the BSC is one of the major innovations of recent decades in management systems. However, notwithstanding the progress achieved, the BSC is not without its own shortcomings. As reported in the literature (*e.g.* Ittner *et al.* 1997; Ferreira *et al.* 2011), it seems generally agreed that the way compensations between criteria within a BSC framework are calculated remains an open issue. Following this, and considering the versatility and great potential of the multiple criteria decision analysis (MCDA) techniques (*e.g.* MACBETH) in dealing with trade-offs among evaluation criteria, this paper aims to analyze possible contributions of the MCDA approach to overcome the measurement shortcomings of the BSC. It must be highlighted, however, that Zorzi and Ensslin (2006) have already reported the integrated use of the BSC and MACBETH in the construction of a performance measurement system in an accounting context. As such, this paper builds on previous work, but aims to discuss new findings and amplify the interest of the MCDA approach to other measurement contexts in order to broaden the generalizability of the results.

The paper is organized as follows. Section 2 presents the BSC as a generic method of performance evaluation, and highlights its major shortcomings. Section 3 presents concepts related to MCDA, respective background and potential in the cardinal measurement of trade-offs among evaluation criteria. In section 4, a numerical example is provided. Section 5 discusses the pros and cons associated with the integrated use of the BSC and MACBETH, and concludes the paper.

2. The Balanced Scorecard

The Balanced Scorecard (BSC), as a generic method for performance evaluation, was created and developed by Robert Kaplan and David Norton in the early 1990s (Kaplan, Norton 1992). One of the main reasons for its creation is directly related to the fact that all the classic financial indicators are insufficient to measure the creation of value by the intangible assets of an organization. However, according to Kaplan and Norton (1996a), *"the scorecard wasn't a replacement for financial measures; it was their complement"*. Following this, the BSC has been emphasized in several studies as a tool to assess performance, because it articulates several indicators, while considering the organization's strategy.

The structure of the model is focused on the initial strategic options defined by the organization. From this point of view, the evaluation is done according to a complex set of causes and effects articulated to define the so-called indicators of occurrence (with desirable well-defined goals and well-measured end-points). The modeling system is then initiated by defining key performance variables in the financial area, such as: *return on invested capital* and *sales growth*. Indexed to these types of indicators, different trend indicators (*i.e.* factors that influence the performance of the indicators of occurrence) are identified and defined, which determine the performance of the cen-

tral variables, which are, again, consistent with the strategic goals of the organization. Thus, the BSC emerges from a conceptual basis upon which non-financial aspects can be *"lead indicators of future financial performance"* (Kaplan, Norton 2001a). That is, the process is deployed in sequence for the dimensions of innovation and learning, internal processes and customers, determining the financial perspective (*i.e.* innovation and learning influence internal processes, which influence the customers' perspective; finally, the customer's perspective influences the financial perspective).

Despite the fact that the BSC has been criticized on the grounds of being too linear in the analysis of the causal relationships among evaluation criteria, the model has considerable merit, depending on its contribution to the systematization of the dynamic action of organizations. As a result, one may catalog the BSC as a tool for strategic management integration (Kaplan, Norton 1996a, 1996b).

2.1. Explanation and scope of the method

As previously stated, the BSC is a tool that provides a dynamic view of the organization, targeting the financial, customer, internal processes, and learning and growth perspectives, all aligned with the organization's strategy, which must be clearly known and correct. *Figure 1* presents the conceptual scheme of the BSC.
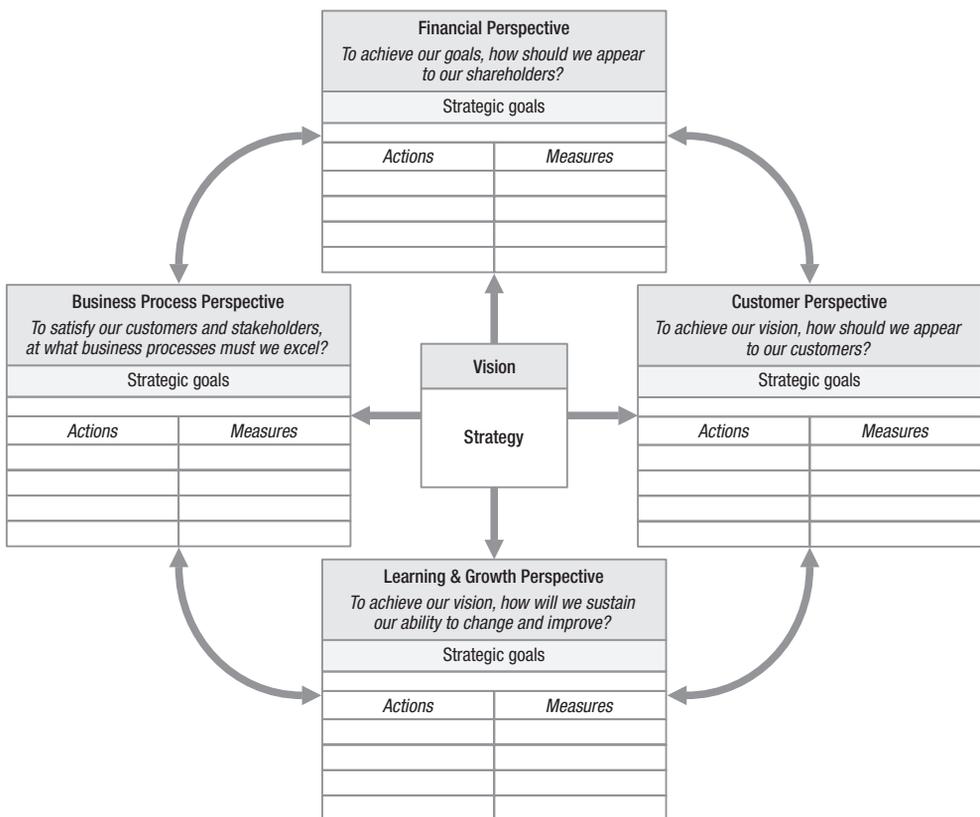

Fig. 1. Conceptual scheme of the BSC
Source: Adapted from Kaplan and Norton (1996a).

In practical terms, the model reflects an organization's vision and strategy, seeking a balance between the various external indicators (*i.e.* shareholder-oriented or customer-oriented) and internal measures of critical business processes (such as innovation, learning and growth). Consequently, the BSC seeks to obtain information on the segment in which the organization competes, conquers clients, creates value, etc. That is, the method aims to offer a future vision and a way to achieve it (Kaplan, Norton 2001b). Nevertheless, it is noted that the BSC was not designed as a tool of quantification, although it could (or should) be used by executives who need to make decisions regarding operations, production processes, objectives, products and customers. Thus, the BSC is based on four perspectives, which when integrated and analyzed together, provide a balanced view of the current and future positions of business performance, producing a balance between: (1) short-term objectives and long-term financial and non-financial indicators; (2) externally oriented assessments (shareholders or customers) and internally oriented evaluations (internal business processes, innovation and learning); and (3) evaluations of past efforts and drivers of future performance. Furthermore, as stated by Kaplan and Norton (1996a), the BSC is also useful for communicating the mission that exists throughout the organization.

2.1.1. Financial perspective

Fleisher and Mahaffi (1997) argue that *"no company survives without ensuring its ongoing financial viability"*. Sharing this view, the BSC aims to evaluate the financial results of a strategy. That is, the method aims to evaluate business growth and development, as well as the shareholders' satisfaction. Among the diversified financial indicators that can be considered, it is typical to include: *return on investment*, *economic added value*, *profitability*, *growth revenues*, *costs reduction* and other financial goals that should be aligned with the company's strategy.

2.1.2. Customer perspective

From the customer perspective, the method aims to target market segments and identify measures of success to evaluate those same segments. In broad terms, customers' concerns are usually supported on four variables: time, quality, service performance and cost. In terms of indicators considered essential, one may highlight: *market share*, *customer's capability of acquisition*, *customer retention*, *profitability* and *customer's level of satisfaction*, among others (Kaplan, Norton 1992, 2001a).

2.1.3. Internal process perspective

The internal process perspective is analyzed according to guidelines provided by the financial and customer perspectives. The internal processes are the various activities undertaken within the organization, ranging from identification of needs to customers' satisfaction. This view encompasses the innovation process (creation of products and services), operations (production and marketing), and after-sales services (customer support after the sale). The improvement of internal processes is a key indicator of future financial success. For the indicators, it is typical to consider: *number of innovations* and *number of operations after the sale*.

2.1.4. Learning and growth perspective

The learning and growth perspective offers the basis for achieving the objectives of the other three perspectives. To this end, it identifies the infrastructure needed to promote growth and long-term improvements, which comes from three main sources: people, systems and organizational procedures. At this stage, it also identifies the capabilities that a company should have in terms of internal processes to be able to create value for customers and shareholders. As indicators, it is typical to consider: *the level of employee satisfaction*, *employee turnover*, *profitability per employee* and *employee engagement with suggestions that facilitate improvements*.

2.2. General and specific limitations of the BSC applications

Following the earlier discussion, it seems clear that the BSC is a robust method that adds value to organizations by providing relevant information, creating environmental conditions conducive to individual and collective (through learning and participation) improvement, and eliminating the need to apply a specific control system for each situation. In fact, as stated by Kaplan and Norton (1992), *"the scorecard puts strategy and vision, not control, at the center"*. In this sense, the adoption of an evaluation system based on the BSC requires a clear definition of the strategy, as well as a team willing to invest effort in an initiative that is often difficult to quantify. Moreover, according to Ottoboni *et al.* (2002), this method still lacks understanding and simplification to be properly applied to many potential scenarios. Consequently, some authors (*e.g.* Fleisher, Mahaffi 1997; Ottoboni *et al.* 2002; Curado, Manica 2010) have been pointing out several limitations to the BSC, such as: (1) it preserves the emphasis on financial results by submitting all measures to financial objectives; (2) its practical implementation *"takes a considerable amount of time (up to 30 months) to design and implement as a strategic management system"* (Fleisher, Mahaffi 1997); (3) its formulation may depend on the relative negotiation power of different groups; (4) other stakeholders, not previously considered in the four perspectives, may arise; (5) it does not explore the interrelationships between partial indicators and the organization's overall performance, such that it is possible to observe, simultaneously, a set of values presenting satisfaction and other set of values presenting dissatisfaction; (6) the analysis may be limited by "interpretation effects"; and (7) raises doubts on the way that relative weights are chosen to balance the evaluation criteria (for a categorized list of BSC limitations, see Curado, Manica 2010).

Despite the relative specificity of these limitations, it seems clear that from their combination two main lines of criticism directed toward the BSC emerge. On the one hand, and according to Fleisher and Mahaffi (1997), *"the biggest problem practitioners will face in adopting a balanced scorecard approach […] will be in developing measures and installing the information systems needed to capture the data underlying the measures"*. However, Kaplan and Norton (2000) propose the use of strategic maps to overcome this particular limitation. On the other hand, and despite recent progress, it seems generally agreed that the way compensations between criteria within a BSC framework are calculated *"remains an open issue"* (Ittner *et al.* 1997; Ferreira *et al.* 2011). Following this, one of the contributions of this study is, precisely, to propose the use of mul-

tiple criteria decision analysis (MCDA) methods to bring simplicity and transparency to the calculation of compensations (*i.e.* trade-offs) among evaluation criteria within a BSC framework. In the next section, the MCDA background is briefly presented.

3. Brief presentation of the MCDA background

As far as the treatment of complex problems is concerned, MCDA can be seen as a new Operational Research (OR) branch. As defended by Bana e Costa *et al.* (1997) *"in contrast to the more classical OR approaches, the multicriteria decision aid framework facilitates learning about the problem and the alternative courses of action, by enabling people to think about their values and preferences from several points of view"*. Thus, by recognizing the limits of the objective approaches and by considering that decision makers should shape and/or transform their own preferences based on their own judgments of value, multicriteria methods aim to construct something that does not pre-exist. Keeney (1992) strengthens this line of thinking, affirming that *"values are subjective, but they undeniably are part of decision situations. Not modeling them does not make them go away"*. As depicted in *Figure 2*, in dealing with complex problems, the need to consider subjective aspects becomes evident.
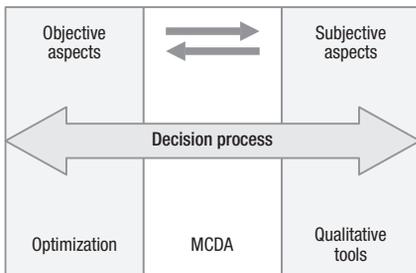

Fig. 2. MCDA conceptual approach

As can be seen through the conceptual scheme presented in *Figure 2*, MCDA highlights the existing inseparability between the objective aspects and the subjective aspects of a decision-making process. Moreover, through an interactive and constructive supporting decision process, MCDA provides decision makers with arguments that enable them to reflect, to readjust and/or to validate their own convictions and judgments of value. Quoting Bana e Costa *et al.* (1997), *"the theory of MCDA is thus an open theoretical field and not a closed mathematical theory solving a specific class of problems"*. Roy and Vanderpooten (1997) discuss some of the features that characterize the MCDA approach (*Table 1*).

Table 1. Main features of the MCDA approach

- Depending on the variables defined, the boundary between what is feasible and what is not feasible is vague and frequently changes.
- In many real-life problems, the decision maker is either difficult to identify or simply does not exist. Rather, s/he is the person (or group of people) on behalf of whom the support should be provided.
- Information is often vague, ill-determined or uncertain.
- The study itself serves to resolve conflicts and/or contradictions among the actors involved in the decision-making process.
- Usually, it seems unreasonable to say if a decision is good or bad by purely referring to a mathematical model. Often the issues that matter involve dimensions of different nature, which also contribute to the quality/success of the final decision.

Source: Adapted from Roy and Vanderpooten (1997).

The features included in *Table 1* show that objective factors and alternatives' characteristics interact with factors of a subjective nature. Therefore, to omit the importance of subjective factors or simply putting them aside in the decision-making process seems to be a negligent thing to do. As it is widely discussed in the MCDA literature, the main gap of any single criterion approach is related to the non-recognition that objectivity has limits. In fact, the search for optimization is emphasized even when the existence of multiple objectives is considered. In line with Belton and Stewart (2002), *"the concept of an optimum does not exist in a multicriteria framework and thus multicriteria analysis cannot be justified within the optimisation paradigm frequently adopted in traditional OR [...]. MCDA is an aid to decision-making, a process which seeks to: integrate objective measurement with value judgments; make explicit and manage subjectivity"*.

As can be observed, this criticism to the single criterion methodologies seems to be supported on the basis that subjectivity is inherent in all decision-making processes and, therefore, one of the major aims of the MCDA approach is to make individuals' preferences explicit, while ensuring transparency in the decision-making process. As defended by Belton and Stewart (2002) *"[...] every decision [...] requires the balancing of multiple factors [...] sometimes explicitly, sometimes without conscious thought"*. In this regard, from a constructivist perspective, the MCDA approach recognizes the limits of the mathematical optimum, and defends a guiding principle that, without preconditions, models are supported on the observation of working hypotheses and/or on a set of key elements.

3.1. Measuring trade-offs among criteria

As has been discussed herein, the way compensations between criteria within a BSC framework are calculated *"remains an open issue"* (Ittner *et al.* 1997; Ferreira *et al.* 2011). Nonetheless, due to their remarkable heuristic elasticity that provides decision makers with a bigger capacity to support decisions and to adapt strategies to a permanent changing environment, MCDA methodologies seem to be suitable to overcome this shortcoming of the BSC. In fact, the assessment of weights to evaluation criteria has been discussed in the MCDA literature for a long time (Goodwin, Wright 1991; Junior 2008). In the next subsection, one of the most well-known MCDA techniques is presented, in order to exemplify the way trade-offs among criteria can be obtained.

3.2. The MACBETH technique

MACBETH stands for Measuring Attractiveness by a Categorical Based Evaluation Technique. It was created during the 1990s by Carlos Bana e Costa and Jean Claude Vansnick (Bana e Costa, Vansnick 1994, 1997, 1999) and, in broad terms, is an interactive technical procedure designed for supporting the construction of numerical scales of intervals, which aim to quantify the difference of attractiveness between elements of a certain set, throughout a learning process and supported by visual interactive software (*i.e.* M-MACBETH).

The technique uses a simple question-answer procedure, which consists of asking decision makers to pairwise compare options by giving a qualitative judgment of the difference in attractiveness between them. Following a constructivist approach, the procedure

also tests the consistency of the answers, offers suggestions to bypass inconsistent situations and provides the analysis for aiding the decision maker to enter the domain of cardinal measurement (for a general overview and some practical applications, see Bana e Costa, Vansnick 1994, 1997, 1999; Bana e Costa *et al.* 2005). Thus, it seems to be an extremely useful technique not only in the construction of cardinal value functions, but also in supporting the definition of trade-offs among evaluation criteria (*Figure 3*).

Unlike the technical procedure used in the direct rating and bisection, in which decision maker/s is/are subject to a cognitive effort that involves more than two actions (*e.g.* is the difference of attractiveness between *a* and *b* greater, smaller or equal to the difference of attractiveness between *c* and *d*?), in the MACBETH methodology, the process involves only two actions at a time. Therefore, Bana e Costa and Vasnick (1994) emphasize that, as in the Analytic Hierarchy Process (AHP) of Saaty (1980), the process should be simple and natural, involving only the drafting of judgments between each pair of actions. The novelty lies, however, in the introduction of a semantic scale formed by categories of difference of attractiveness.

In practice, as is acknowledged by the authors, this type of approach had been previously advanced by Freeling and Belton (Bana e Costa, Vasnick 1994) in the context of criticisms made to the AHP (Boucher, MacStravic 1991; Weber 1993; Davies 1994; Bana e Costa *et al.* 1997; Belton, Stewart 2002; Dubois 2003; Bana e Costa, Vansnick
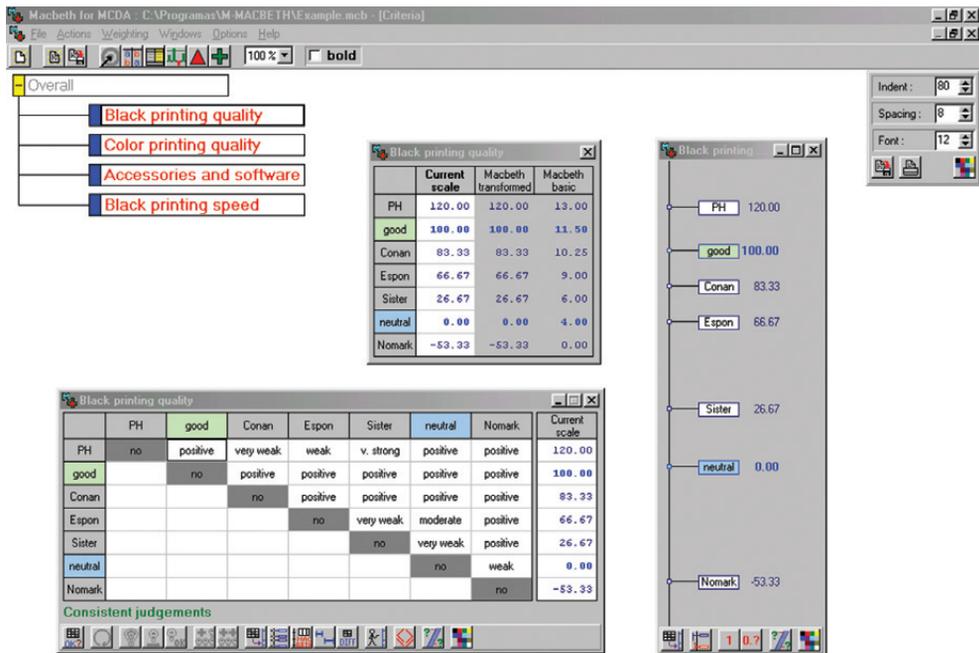

Fig. 3. The *MACBETH* approach (Matrix of judgments, weights and cardinal value scale)

2008). However, the authors argue that the MACBETH approach addresses the issue of cardinal value scales in a more innovative way because, unlike the AHP, which uses scales of ratio of priority or importance (for AHP applications see *e.g.* Korsakienė 2004; Podvezko 2009), the MACBETH methodology makes use of scales of differences of attractiveness, which are nothing more than interval scales (Bana e Costa, Vansnick 1994, 2008). As such, it seems clear that the concept of attractiveness takes a different meaning from the concepts of priority or importance, since it allows negative values representing repulsive feelings, while the latter two (importance and priority) do not fall below the zero level. Following this, we can assume that the main disadvantage of working with mono-polar concepts occurs precisely because zero does not translate to a neutral level, but at a clearly negative level. In addition, there might be barriers related to lack of substantive significance, particularly if the answers to the questions are interpreted in terms of intensity of preference.

In line with Bana e Costa and Vansnick (1994), the initial framework of the MACBETH methodology is anchored to the problem of numerical representation of semi-orders for multiple thresholds, resolved by Doignon. According to the authors, in a problem where there is a structure of m binary relations $[P^{(1)}, ... , P^{(k)}, ... , P^{(m)}]$, where $P^{(k)}$ stands for a preference that is stronger the greater the k, based on a certain point of view PV_j, the conversion of these relationships of preference in numbers becomes possible. In fact, the numerical codification procedure proposed by the MACBETH methodology consists in associating to each action of X (with $X = \{a, b, ... , n\}$ being a finite set of n actions), a value x (resulting from $v(.): X \rightarrow R$) such that differences as $v(a) - v(b)$ (with a more attractive than b (*i.e. a P b*)), are as compatible as possible with the value judgments made by the decision maker/s. In other words, for all pairs (a, b) assigned to a certain category of difference of attractiveness C, the differences $v(a) - v(b)$ will belong to the same interval (without overlaps) (Bana e Costa, Vansnick 1994)). Whereas two contiguous ranges correspond to two consecutive categories, the technical procedure is to associate asymmetric partitions of the ray of positive reals to partition classes of ordered pairs (a, b) (with $a P b$) (*Figure 4*).

Following this, and in order to define the intervals, we just need to set the limits, which can be understood as transition thresholds between categories of consecutive differences of attractiveness. In this sequence, semi-multiple orders are naturally introduced as long as we wish to represent preferences through a value function v and thresholds of functions s_k, such as:

$$a P^{(k)} b: \quad s_k < v(a) - v(b) < s_{k+1}. \tag{1}$$

It is worth recalling that the thresholds s_k are positive real constants, which permit definition of the intervals corresponding to the semantic differences of attractiveness. Accordingly, a range of differences of attractiveness has to be limited on its left by "its" zero, but are not limited on its right. As stated by Bana e Costa and Vansnick (1994), being $a P^{(m)} b$, it is always theoretically possible to find a level of
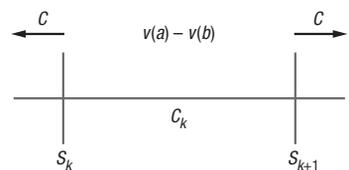

Fig. 4. Allocation of $v(a) - v(b)$ to a category C_k

impact by setting an actual or fictitious action c, such that c is more attractive than b, more than a is more attractive than b. Following this logic, the final semantic category C_m cannot be limited on the right and, between the origin $s_1 = 0$ and s_m, an infinite number of categories and thresholds can be defined. *Figure 5* illustrates an example of a range of categories of difference of attractiveness.
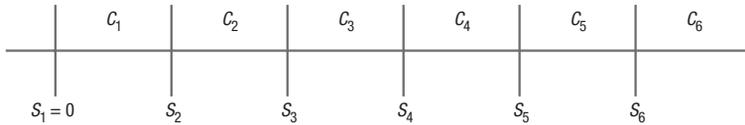

Fig. 5. Scale of categories of difference of attractiveness

As defended by Bana e Costa *et al.* (2005), *"the basic idea underlying the initial development of MACBETH was that limits of these intervals should not be arbitrarily fixed a priori, but determined simultaneously with numerical value scores for the elements of X"*. Therefore, and from a practical point of view, the methodology starts by asking the decision-maker/s to allocate the difference of attractiveness between each pair of actions $(a, b) \in X$ to one of the following semantic categories: C_0 = Null (or indifference (*i.e. a* I *b*)); C_1 = Very weak; C_2 = Weak; C_3 = Moderate; C_4 = Strong; C_5 = Very strong; and C_6 = Extreme (Bana e Costa *et al.* 2005). For example, if the decision maker considers a more attractive than b and the difference between both actions is strong, then $(a, b) \in C_4$. Based on the value judgments of the decision-maker/s, the methodology will propose, if possible, an initial scale that satisfies formulations (2) and (3) (Junior 2008).

$$\forall a, b \in X : v(a) > v(b) \Leftrightarrow aPb, \tag{2}$$

$$\forall k, k^* \in \{1, 2, 3, 4, 5, 6\}, \forall a, b, c, d \in X \text{ with } (a, b) \in C_k$$
$$\text{and } (c, d) \in C_{k*} : k \geq k^* + 1 \Rightarrow v(a) - v(b) \geq v(c) - v(d). \tag{3}$$

Linear programming is then applied as follows (Junior 2008):

> Min $v(n)$
> $S.T. : \forall a, b \in X : aPb \Rightarrow v(a) \geq v(b) + 1$
> $\quad \forall a, b \in X : aIb \Rightarrow v(a) = v(b)$
> $\quad \forall (a, b), (c, d) \in P$, if the difference of attractiveness between
> $\quad\quad\quad a$ and b is bigger than between c and d, then:
> $\quad\quad\quad v(a) - v(b) \geq v(c) - v(d) + 1 + \delta(a, b, c, d)$
> $v(a^-) = 0$

where:

n is an element of X so that $\forall a, b, c, ... \in X : n(P \cup I) a, b, c, ...$

a^- is an element of X so that $\forall a, b, c, ... \in X : a, b, c, ... (P \cup I) a^-$

$\delta(a, b, c, d)$ is the minimal number of categories of difference of attractiveness
$\quad\quad$ between the difference of attractiveness between a and b and the
$\quad\quad$ difference of attractiveness between c and d. $\tag{4}$

Among other developments, recent progress on the MACBETH framework has allowed for implementation of an automatic procedure for detecting inconsistencies, even in contexts of incomplete arrays of value judgments. From a practical standpoint, substantial improvements in the development of this approach were also achieved by increasing the capabilities of this innovative methodology for decision support. It is important to note, however, that the MACBETH approach is not without its own limitations. In fact, the technique requires an enormous willingness on the part of decision makers, and a high dedication on the part of the facilitator. On the other hand, depending on the number of levels of impact defined for each descriptor, completing the matrices can become a demanding task for the actors involved in the process and, as such, difficulties in gathering data may arise (for further details, see Ferreira *et al.* 2011). Nonetheless, following Bana e Costa *et al.* (2003), *"its essential characteristics, however, have never changed; in fact MACBETH has always remained consistent with the ideas that led to its creation as an humanistic, interactive and constructive approach to the problem of how to build a quantitative model of values based on qualitative (verbal) difference judgements, that facilitates the path from ordinal to cardinal preference modelling"*. It is precisely by bringing together the humanistic, interactive and constructivist strands that the MACBETH approach is characterized as a multiple criteria decision analysis technique with great potential in the context of measurement.

4. A numerical example

In order to provide a numerical example of the potential of MCDA in the context of cardinal measurement of trade-offs, let us consider two hypothetical criteria CRT_1 and CRT_2, which could be two of the variables included (or to be included) in one of the four scorecards (*i.e.* perspectives) presented in *Figure 1*. From a simplistic and merely exemplificative point of view, the process may start by asking the decision maker (or group of decision makers within a negotiation process) to rank criteria according to their degree of attractiveness (*i.e.* if the decision maker considers CRT_1 more attractive than CRT_2, then CRT_1 should be ranked in first place). The next step consists in eliciting from the decision maker qualitative judgments regarding the difference of attractiveness between the criteria. This is made based on the semantic categories of value judgments presented in *subsection 3.2*. As each judgment is given, the M-MACBETH software automatically verifies the matrix's consistency and suggests modifications in case of any

detected inconsistency. *Figure 6* depicts an example of a consistent matrix, where CRT_1 and CRT_2 are evaluated based on the difference of attractiveness between them.

As can be observed in *Figure 6*, a numerical scale is calculated from the complete and consistent matrix of judgments. In this merely exemplificative case, 54.55% and 45.45% are, respectively, the weights of CRT_1 and CRT_2 (*Figure 7*).
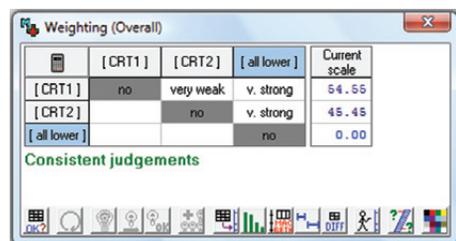

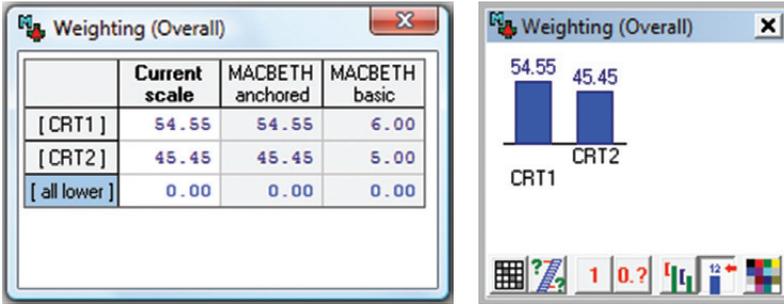**Fig. 6.** Matrix of qualitative judgments

Fig. 7. Scale and weights

The scale should always be discussed to ensure that it adequately represents the magnitude of the decision maker's judgments. In case of disagreement, the weights should be adjusted, respecting the limits proposed, to maintain the consistency of the previously given answers. Obviously, sensitivity, robustness and dominance analyses should be carried out in order to allow the decision maker to see how a change in any of the weights would affect the overall results. Following this, one may state that MACBETH offers a very good mathematical basis, and its friendliness in obtaining trade-offs among evaluation criteria should be highlighted.

5. Conclusions and future research

As previously stated, the current economic thinking assumes that performance evaluation is a crucial element for the promotion of improvement initiatives. From this perspective, remarkable progress has occurred over the past two decades in the development of performance measurement frameworks. Among the most well known examples of those frameworks, the BSC should be highlighted. However, despite the progress achieved by the BSC, the framework is not without its own limitations.

Different types of criticism have been directed toward the BSC, but it should be underlined that there appears to be little concern about the use of structured approaches to guide managers in explicitly dealing with trade-offs, helping them to prioritize performance objectives. As such, possible contributions from the MCDA research field have been discussed in this paper, namely the use of the MACBETH technique in the calculation of compensations among evaluation criteria. As clearly indicated, this paper is not empirical research and builds on previous work. Nonetheless, this study aimed to amplify the discussion and interest of the MCDA approach to other measurement frameworks in order to increase the generalizability of the results achieved by Zorzi and Ensslin (2006) in an accounting context. By underlining the good mathematical basis and user friendliness of the MACBETH approach, we believe that the development and promotion of the integrated use of MCDA tools in BSC applications to measure trade-offs among evaluation criteria in the BSC framework can be a key factor in ensuring more successful implementations and uses of the BSC within firms. As such, further research is strongly encouraged and, among other things, it seems relevant to:

(1) determine the robustness of our approach by conducting different case studies in the same industry; (2) increase the framework's generalizability by conducting different case studies within different industries and/or different countries; and (3) increase the reliability of the results creating surveys based on the different case studies' outcomes. Eventual improvements resulting from different case studies in different contexts will contribute to strengthen the robustness, generalizability, reliability and potential of the approach discussed in this paper.

Acknowledgements

References

Acar, A.; Zehir, C. 2010. The harmonized effects of generic srategies and business capabilities on business performance, *Journal of Business Economics and Management* 11(4): 689–711. http://dx.doi.org/10.3846/jbem.2010.34

Bana e Costa, C.; Vansnick, J. 1994. Uma nova abordagem ao problema da construção de uma função de valor cardinal: MACBETH, *Avaliação de Projectos e Decisão Pública*. Fascículo VI: AEIST/UTL.

Bana e Costa, C.; Vansnick, J. 1997. Applications of the MACBETH approach in the framework of an additive aggregation model, *Journal of Multi-Criteria Decision Analysis* 6(2): 107–114. http://dx.doi.org/10.1002/(SICI)1099-1360(199703)6:2<107::AID-MCDA147>3.0.CO;2-1

Bana e Costa, C.; Vansnick, J. 1999. The MACBETH approach: basic ideas, software and an application, in N. Meskens, M. Roubens (Eds.). *Advances in Decision Analysis. Mathematical Modelling: Theory and Applications*, 4. Kluwer Academic Publishers: Dordrecht, 131–157.

Bana e Costa, C.; Vansnick, J. 2008. A critical analysis of the Eigenvalue method used to derive priorities in the AHP, *European Journal of Operational Research* 187(3): 1422–1428. http://dx.doi.org/10.1016/j.ejor.2006.09.022

Bana e Costa, C.; De Corte, J.; Vansnick, J. 2003. MACBETH, *Working Paper* LSE OR 03.56. London: London School of Economics.

Bana e Costa, C.; De Corte, J.; Vansnick, J. 2005. On the mathematical foundations of MAC-BETH, in J. Figueira, S. Greco, M. Ehrgott (Eds.). *Multiple Criteria Decision Analysis: the State of the Art Surveys*. New York: Springer, 409–442.

Bana e Costa, C.; Stewart, T.; Vansnick, J. 1997. Multicriteria decision analysis: some thoughts based on the tutorial and discussion sessions of ESIGMA meetings, *European Journal of Operational Research* 99(1): 28–37. http://dx.doi.org/10.1016/S0377-2217(96)00380-3

Belton, V.; Stewart, T. 2002. *Multiple Criteria Decision Analysis: An Integrated Approach*. Dordrecht: Kluwer Academic Publishers.

Boucher, T.; MacStravic, E. 1991. Multiattribute evaluation within a present value framework and its relation to the analytic hierarchy process, *The Engineering Economist* 37(1): 1–32. http://dx.doi.org/10.1080/00137919108903055

Curado, C.; Manica, J. 2010. Management control systems in Madeira island largest firms: evidence on the Balanced Scorecard usage, *Journal of Business Economics and Management* 11(4): 652–670. http://dx.doi.org/10.3846/jbem.2010.32

Davies, M. 1994. A multicriteria decision model application for managing group decisions, *Journal of the Operational Research Society* 45(1): 47–58. http://dx.doi.org/10.1057/jors.1994.6

Dubois, D. 2003. Evaluation and decision models: a critical perspective (book review), *Fuzzy Sets and Systems* 139(2): 469–472. http://dx.doi.org/10.1016/S0165-0114(03)00038-1

Ferreira, F.; Santos, S.; Rodrigues, P. 2011. Adding value to bank branch performance evaluation using cognitive maps and MCDA: a case study, *Journal of the Operational Research Society* 62(7): 1320–1333. http://dx.doi.org/10.1057/jors.2010.111

Fleisher, C.; Mahaffi, D. 1997. A Balanced Scorecard approach to public relations management assessment, *Public Relations Review* 23(2): 117–142. http://dx.doi.org/10.1016/S0363-8111(97)90020-5

Goodwin, P.; Wright, G. 1991. *Decision Analysis for Management Judgment*. John Wiley.

Ittner, C.; Larcker, D.; Meyer, M. 1997. *Performance, Compensation, and the Balanced Scorecard*. The Wharton School, The University of Pennsylvania.

Junior, H. 2008. Multicriteria approach to data envelopment analysis, *Pesquisa Operacional* 28(2): 231–242. http://dx.doi.org/10.1590/S0101-74382008000200004

Kaplan, R.; Norton, D. 1992. The Balanced Scorecard: measures that drive performance, *Harvard Business Review* 70(1): 71–79.

Kaplan, R.; Norton, D. 1996a. Using the Balanced Scorecard as a strategic management system, *Harvard Business Review* Jan/Feb: 75–85.

Kaplan, R.; Norton, D. 1996b. Strategic learning and the Balanced Scorecard, *Strategy & Leadership* 24(5): 18–24. http://dx.doi.org/10.1108/eb054566

Kaplan, R.; Norton, D. 2000. Having trouble with your strategy? Then map it, *Harvard Business Review* 78(5): 167–176.

Kaplan, R.; Norton, D. 2001a. Transforming the Balanced Scorecard from performance measurement to strategic management: part I, *Accounting Horizons* 15(1): 87–104. http://dx.doi.org/10.2308/acch.2001.15.1.87

Kaplan, R.; Norton, D. 2001b. Transforming the Balanced Scorecard from performance measurement to strategic management: part II, *Accounting Horizons* 15(2): 147–160. http://dx.doi.org/10.2308/acch.2001.15.2.147

Keeney, R. 1992. *Value-Focused Thinking: A Path to Creative Decision Making*. Harvard: Harvard University Press.

Korsakienė, R. 2004. Determining competitive advantage: the analytic hierarchy process, *Journal of Business Economics and Management* 5(4): 205–215. http://dx.doi.org/10.1080/16111699.2004.9636084

Ottoboni, C.; Fernandes, C.; Pamplona, E.; Pagni, T. 2002. Algumas Razões para a Ocorrência de Falhas na Implementação do Balanced Scorecard (BSC), in *Proceedings of the XXII ENEGEP*, Curitiba, Brazil, 23–25 October, 1–8.

Podvezko, V. 2009. Application of AHP technique, *Journal of Business Economics and Management* 10(2): 181–189. http://dx.doi.org/10.3846/1611-1699.2009.10.181-189

Roy, B.; Vanderpooten, D. 1997. An overview on "The European School of MCDA: Emergence, Basic Features and Current Works", *European Journal of Operational Research* 99(1): 26–27. http://dx.doi.org/10.1016/S0377-2217(96)00379-7

Saaty, T. 1980. *The Analytic Hierarchy Process*. New York: McGraw-Hill.

Santos, S.; Belton, V.; Howick, S. 2008. Enhanced performance measuring using OR: a case study, *Journal of the Operational Research Society* 59(6): 762–775. http://dx.doi.org/10.1057/palgrave.jors.2602397

Strandskov, J. 2006. Sources of competitive advantage and business performance, *Journal of Business Economics and Management* 7(3): 119–129.
http://dx.doi.org/10.1080/16111699.2006.9636132

Urbonavičius, S.; Ivanauskas, R. 2005. Evaluation of multiple retailers' market positions on the basis of image and attributes measurement, *Journal of Business Economics and Management* 6(4): 199–206. http://dx.doi.org/10.1080/16111699.2005.9636109

Weber, S. 1993. A modified analytic hierarchy process for automated manufacturing decisions, *Interfaces* 23(4): 75–84. http://dx.doi.org/10.1287/inte.23.4.75

Zinkevičiūte, V. 2007. Evaluation of business strategic decisions under changing environment conditions, *Journal of Business Economics and Management* 8(4): 267–274.
http://dx.doi.org/10.1080/16111699.2007.9636179

Zorzi, A.; Ensslin, S. 2006. Interação entre o BSC e a Metodologia MCDA-C: Construção de um Modelo de Gestão para o Setor de Contabilidade de uma Entidade Fechada de Previdência Complementar, in *Proceedings of the XIII SIMPEP*, Bauru/São Paulo, Brazil, 6–8 November, 1–12.

Fernando A. F. FERREIRA is an Adjunct Professor and Vice-President of the Scientific Council at the School of Management and Technology of the Polytechnic Institute of Santarém, Portugal, and Adjunct Research Professor at the University of Memphis, TN, USA. He holds a PhD in Quantitative Methods Applied to Economics and Management (2008) from the Faculty of Economics of the University of Algarve, Portugal. In 2009, he was granted with a *Calouste Gulbenkian Foundation* fellowship for Post-Doctoral studies at the Fogelman College of Business and Economics of the University of Memphis. He has authored books, book chapters, peer-reviewed papers and made many conference presentations. He is Associate Editor of the *Journal of Information and Operations Management* (JIOM), and Editorial Board Member of the *International Journal of Innovation and Learning* (IJIL), *International Journal of Intercultural Information Management* (IJIIM), *International Journal of Information and Operations Management Education* (IJIOME) and *International Journal of Learning and Change* (IJLC). His main research interests include multiple criteria decision analysis and integrated systems for performance measurement in the banking sector.